arXiv:1506.00448v1 [math.CO] 1 Jun 2015

A STRUCTURE THEOREM FOR SETS OF SMALL POPULAR DOUBLING, REVISITED

PRZEMYSŁAW MAZUR

ABSTRACT. We prove that every set $A \subset \mathbb{Z}/p\mathbb{Z}$ with $\mathbb{E}_x \min(1_A * 1_A(x), t) \leqslant (2 + \delta)t\mathbb{E}_x 1_A(a)$ is very close to an arithmetic progression. Here p stands for a large prime and δ, t are small real numbers. This shows that the Vosper theorem is stable in the case of a single set.

1. INTRODUCTION

In the recent paper [Maz15] we proved a structure theorem for sets of integers having small popular doubling. We were aiming to extend this theorem to make it also work for sets of residue classes modulo a prime. Unfortunately we were unable to achieve this using the methods of that paper. In this paper we prove that missing statement using entirely different methods. To be more specific, our goal is to prove the following statement.

Theorem 1.1. *Let $0 < \alpha_1 < \alpha_2 < \frac{1}{4}$ and $\eta > 0$. Then there exist positive real numbers $\delta_0 = \delta_0(\alpha_1, \alpha_2, \eta)$, $C = C(\alpha_1, \alpha_2, \eta)$ and $p_0 = p_0(\alpha_1, \alpha_2, \eta)$ with the following properties. Let $p > p_0$ be a prime and let $A \subset \mathbb{Z}/p\mathbb{Z}$ be a set. Suppose that the density $\alpha = \frac{|A|}{p}$ satisfies $\alpha_1 < \alpha < \alpha_2$. Furthermore, suppose that*

$$\mathbb{E}_x \min(1_A * 1_A(x), t) \leqslant (2 + \delta)\alpha t$$

for some numbers $\delta \in (0, \delta_0)$ and $t \in (0, t_0(\alpha_1, \alpha_2, \eta, \delta))$. Then there is an arithmetic progression P with $|P| \leqslant (1 + (1 + \eta)\delta)\alpha p$ and $|A \setminus P| \leqslant C(\delta\alpha)^{1/2}p$.

To fix the notation, let us use the Haar probability measure on all groups appearing in this paper. That means that the symbol $\mathbb{E}_x$ used above is just a shorthand for $\frac{1}{p}\sum_x$, and by $f * g(x)$ we mean $\mathbb{E}_y f(y)g(x - y)$.

The dependences in the statement of Theorem 1.1 look rather complicated, let us justify them a little bit. If we were dealing with sets satisfying just $|A + A| \leqslant (2 + \delta)\alpha p$, then the correct bound for the size of P would be $|P| = |A + A| - |A| + 1 = (1 + \delta)\alpha p + 1$ (see [SZ09] for details). The parameter η indicates that we can make as small error as we like, even in terms of δ, but to achieve that the popularity parameter t has to be sufficiently small, in terms of both η and δ. Ideally we would like to conclude that $|P| \leqslant ((1 + \delta)\alpha + O(t))p$ (see [Maz15]), but with our methods we are unable to achieve that.

2. PROOF OF THE MAIN RESULT

Let us start with fixing all the parameters and the set $A \subset \mathbb{Z}/p\mathbb{Z}$ satisfying the assumption. We intend to apply the arithmetic regularity lemma (Theorem A.9) to the function $f = 1_A$. Let $\varepsilon > 0$ and $\mathcal{F}$ be a growth function to be specified later. Then we can write $f = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$, as in the statement of Theorem A.9. Let us first get rid of the function f_{unf}.

Lemma 2.1. *Let $g, h : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ be functions. Then the following inequality holds:*

$$\|g * h\|_2 \leqslant \|g\|_{U^2}\|h\|_{U^2}$$

1

Proof. Using Parseval's identity and the relation between the convolution and Fourier transform we get

$$\|g * h\|_2^2 = \sum_r |\widehat{g * h}(r)|^2 = \sum_r |\widehat{g}(r)|^2 |\widehat{h}(r)|^2.$$

On the other hand, we know that

$$\|g\|_{U^2}^4 = \sum_r |\widehat{g}(r)|^4 \qquad \text{and} \qquad \|h\|_{U^2}^4 = \sum_r |\widehat{h}(r)|^4.$$

The inequality is then equivalent to the Cauchy-Schwarz inequality in the following form:

$$\left(\sum_r |\widehat{g}(r)|^2 |\widehat{h}(r)|^2 \right)^2 \leqslant \left(\sum_r |\widehat{g}(r)|^4 \right) \left(\sum_r |\widehat{h}(r)|^4 \right).$$

$\square$

Corollary 2.2. Let $1_A = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$ as above. Then the following inequality holds:

$$\mathbb{E}_x \min((f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}}), t) \leqslant (2 + \delta)\alpha t + \frac{2}{(\mathcal{F}(M))^{1/2}}.$$

Proof. First of all, since for all characters χ we have $|\widehat{f_{\mathrm{unf}}}(\chi)| = |\langle f_{\mathrm{unf}}, \chi \rangle| \leqslant \frac{1}{\mathcal{F}(M)}$, we can estimate the U^2 norm of f_{unf} as

$$\|f_{\mathrm{unf}}\|_{U^2}^4 = \sum_\chi |\widehat{f_{\mathrm{unf}}}(\chi)|^4 \leqslant \frac{1}{(\mathcal{F}(M))^2} \sum_\chi |\widehat{f_{\mathrm{unf}}}(\chi)|^2 =$$

$$= \frac{\|f_{\mathrm{unf}}\|_2^2}{(\mathcal{F}(M))^2} \leqslant \frac{\|f_{\mathrm{unf}}\|_\infty^2}{(\mathcal{F}(M))^2} = \frac{1}{(\mathcal{F}(M))^2}.$$

Therefore for any function $g : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ with $\|g\|_\infty \leqslant 1$ the lemma above gives

$$\|f_{\mathrm{unf}} * g\|_1 \leqslant \|f_{\mathrm{unf}} * g\|_2 \leqslant \|f_{\mathrm{unf}}\|_{U^2} \|g\|_{U^2} \leqslant \|f_{\mathrm{unf}}\|_{U^2} \|g\|_\infty \leqslant \frac{1}{(\mathcal{F}(M))^{1/2}}.$$

Applying this to the functions $g = 1_A$ and $g = f_{\mathrm{str}} + f_{\mathrm{sml}}$ and using triangle inequality we get

$$\|1_A * 1_A - (f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}})\|_1 \leqslant \frac{2}{(\mathcal{F}(M))^{1/2}}.$$

Now we use an easy-to-check inequality $|\min(a, t) - \min(b, t)| \leqslant |a - b|$ for $a = 1_A * 1_A(x)$ and $b = (f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}})(x)$ for any $x \in \mathbb{Z}/p\mathbb{Z}$. Combining them with another instance of triangle inequality yields

$$|\mathbb{E}_x \min(1_A * 1_A(x), t) - \mathbb{E}_x \min((f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}}), t)| \leqslant \frac{2}{(\mathcal{F}(M))^{1/2}},$$

which gives the result. $\square$

We managed to remove f_{unf} from our considerations, now it is time for f_{sml}. To deal with this. let λ be a small quantity to be specified and let $B = \{x \in \mathbb{Z}/p\mathbb{Z} : \|\varphi(x)\| \leqslant \frac{\lambda}{2M}\}$ be a Bohr set. Recall that φ is the homomorphism used to construct f_{str}, for the definition of $\|\varphi(x)\|$, see the appendix. Now let

$$C = \{x \in \mathbb{Z}/p\mathbb{Z} : f_{\mathrm{str}}(x) \geqslant \lambda \quad \text{and} \quad \mathbb{E}_{y \in B} |f_{\mathrm{sml}}(x + y)|^2 \leqslant \varepsilon\}.$$

Intuitively, we take all elements where f_{srt} is somewhat large and where f_{sml} is too small to destroy that. First of all, let us estimate the size of C. The set $C' = \{x \in \mathbb{Z}/p\mathbb{Z} : f_{\mathrm{str}}(x) \geqslant \lambda\}$ has size at least $\sum_x f_{\mathrm{str}}(x) - \lambda p$, since

$$\sum_x f_{\mathrm{str}}(x) = \sum_{x \in C'} f_{\mathrm{str}}(x) + \sum_{x \notin C'} f_{\mathrm{str}}(x) \leqslant |C'| + \lambda p.$$

On the other hand, the set $C'' = \{x \in \mathbb{Z}/p\mathbb{Z} : \mathbb{E}_{y \in B}|f_{\mathrm{sml}}(x+y)|^2 > \varepsilon\}$ has size at most εp, because

$$\varepsilon^2 \geqslant \mathbb{E}_x|f_{\mathrm{sml}}(x)|^2 = \mathbb{E}_x(\mathbb{E}_{y \in B}|f_{\mathrm{sml}}(x+y)|^2) \geqslant \frac{\varepsilon|C''|}{p}.$$

Therefore the size of C can be estimated as $|C| = |C' \setminus C''| \geqslant |C'| - |C''| \geqslant \sum_x f_{\mathrm{str}}(x) - (\lambda + \varepsilon)p$. To make it more explicit, note that from the construction of f_{unf} we see that $\mathbb{E}_x f_{\mathrm{unf}}(x) = 0$. That leads to $\mathbb{E}_x(f_{\mathrm{str}} + f_{\mathrm{sml}})(x) = \mathbb{E}_x 1_A(x) = \alpha$; combining it with $|\mathbb{E}_x f_{\mathrm{sml}}(x)| \leqslant \|f_{\mathrm{sml}}\|_1 \leqslant \|f_{\mathrm{sml}}\|_2 \leqslant \varepsilon$ we get $\mathbb{E}_x f_{\mathrm{str}}(x) \geqslant \alpha - \varepsilon$. In the end it means that $|C| \geqslant (\alpha - 2\varepsilon - \lambda)p$.

Now it is time to see the reason why we defined the set C in this way. To see this, let $x_1, x_2 \in C$ and consider four functions: $f_1, f_2, g_1, g_2 : B \to \mathbb{C}$ defined as:

$$f_i(x) = f_{\mathrm{str}}(x_i + (-1)^i x), \qquad g_i(x) = f_{\mathrm{sml}}(x_i + (-1)^i x).$$

Since $f_{\mathrm{str}} + f_{\mathrm{sml}}$ is a nonnegative function, we have the inequality

$$(f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}})(x_1 + x_2) \geqslant$$
$$\geqslant \mathbb{E}_x(f_{\mathrm{str}} + f_{\mathrm{sml}})(x_1 - x)(f_{\mathrm{str}} + f_{\mathrm{sml}})(x_2 + x)1_B(x) =$$
$$= \frac{|B|}{p}\langle f_1 + g_1, f_2 + g_2 \rangle.$$

Now from the Lipschitz nature of F and the definitions of B and C we know that $f_1(x), f_2(x) \geqslant \frac{\lambda}{2}$ for all $x \in B$, which leads to $\langle f_1, f_2 \rangle \geqslant \frac{\lambda^2}{4}$. Moreover since $\|f_i\|_2 \leqslant \|f_i\|_\infty \leqslant 1$ and $\|g_i\|_2 \leqslant \sqrt{\varepsilon}$ (by the definition of C), we also have $|\langle f_1, g_2 \rangle|, |\langle f_2, g_1 \rangle| \leqslant \sqrt{\varepsilon}$ and $|\langle g_1, g_2 \rangle| \leqslant \varepsilon$. Combining all the inequalities together we get $\langle f_1 + g_1, f_2 + g_2 \rangle \geqslant \frac{\lambda^2}{4} - 2\sqrt{\varepsilon} - \varepsilon$. Also, from the properties of Bohr sets (see for example [TV06]) we know that $\frac{|B|}{p} \geqslant (\frac{\lambda}{2M})^{\dim B} \geqslant (\frac{\lambda}{2M})^M$. Therefore if only $t \leqslant (\frac{\lambda}{2M})^M(\frac{\lambda^2}{4} - 2\sqrt{\varepsilon} - \varepsilon)$, we have just proved that $(f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}})(x) \geqslant t$ for all $x \in C + C$. Since

$$\mathbb{E}_x \min((f_{\mathrm{str}} + f_{\mathrm{sml}}) * (f_{\mathrm{str}} + f_{\mathrm{sml}}), t) \leqslant (2 + \delta)\alpha t + \frac{2}{(\mathcal{F}(M))^{1/2}},$$

we know that in this case we have $|C + C| \leqslant ((2 + \delta)\alpha + \frac{2}{t(\mathcal{F}(M))^{1/2}})p$.

The main term of the above expression is $2\alpha p$, while the main term of the expression bounding the size of C is αp. Therefore if the error terms are sufficiently small, we can make use of Serra-Zémor Theorem (proven in [SZ09]) and conclude that the set C is contained in an arithmetic progression $P \subset \mathbb{Z}/p\mathbb{Z}$ of size $|P| = |C + C| - |C| + 1$. We can assume without loss of generality that $|P| \geqslant |A|$ as we can extend P if necessary. We will come back later to the conditions that must be satisfied, let us now proceed with the proof.

We will examine how the progression P is related to the set A. First of all, since $C \subset P$, we know that there can only be εp elemets x outside P for which $f_{\mathrm{str}} \geqslant \lambda$. Therefore we have the inequality

$$\mathbb{E}_x \max((f_{\mathrm{str}} - 1_P)(x), 0) \leqslant \varepsilon + \lambda.$$

This means that we also have $\mathbb{E}_x \max((1_P - f_{\mathrm{str}})(x), 0) \leqslant \mathbb{E}_x(1_P - f_{\mathrm{str}})(x) + \varepsilon + \lambda$. Adding those two inequalities we get $\|1_P - f_{\mathrm{str}}\|_1 \leqslant \mathbb{E}_x(1_P - f_{\mathrm{str}})(x) + 2(\varepsilon + \lambda)$. The last quantity is then an upper bound for the absolute value of the difference of corresponding Fourier coefficients of 1_P and f_{str}. In other words, $|\langle 1_P - f_{\mathrm{str}}, \chi \rangle| \leqslant \mathbb{E}_x(1_P - f_{\mathrm{str}})(x) + 2(\varepsilon + \lambda)$ for each character χ. On the other hand, we know that

$$|\langle 1_A - f_{\mathrm{str}}, \chi \rangle| \leqslant |\langle f_{\mathrm{sml}}, \chi \rangle| + |\langle f_{\mathrm{unf}}, \chi \rangle| \leqslant \varepsilon + \frac{1}{\mathcal{F}(M)}$$

for each character χ. By triangle inequality it means that for every character χ the following holds:

$$|\langle 1_P - 1_A, \chi \rangle| \leqslant \mathbb{E}_x(1_P - f_{\text{str}})(x) + 3\varepsilon + 2\lambda + \frac{1}{\mathcal{F}(M)} \leqslant$$

$$\leqslant \mathbb{E}_x(1_P - 1_A)(x) + 4\varepsilon + 2\lambda + \frac{1}{\mathcal{F}(M)}.$$

Recall now that P is an arithmetic progression, so one of its (non-trivial) Fourier coefficients is as large as it could possibly be, more precisely there exists χ_1 with

$$|\widehat{1_P}(\chi_1)| = |\langle 1_P, \chi_1 \rangle| = \frac{\sin\left(\frac{(|P|-1)\pi}{p}\right)}{p\sin\left(\frac{\pi}{p}\right)}.$$

Now let z be a unit complex number satisfying $z\widehat{1_P}(\chi_1) = |\widehat{1_P}(\chi)|$. Since $1_P - 1_A = 2 \cdot 1_P - 1_{A \cap P} - 1_{A \cup P}$, we have the following lower bound:

$$|\langle 1_P - 1_A, \chi \rangle| \geqslant \Re(z \cdot \langle 2 \cdot 1_P - 1_{A \cap P} - 1_{A \cup P}, \chi \rangle) \geqslant$$

$$\geqslant \frac{2\sin\left(\frac{(|P|-1)\pi}{p}\right) - \sin\left(\frac{(|A \cap P|-1)\pi}{p}\right) - \sin\left(\frac{(|A \cup P|-1)\pi}{p}\right)}{p\sin\left(\frac{\pi}{p}\right)}.$$

We can rearrange the numerator of the last expression as follows

$$2\sin\left(\frac{(|P|-1)\pi}{p}\right) - \sin\left(\frac{(|A \cap P|-1)\pi}{p}\right) - \sin\left(\frac{(|A \cup P|-1)\pi}{p}\right) =$$

$$= 4\sin\left(\frac{|P \setminus A| \cdot \pi}{2p}\right)\sin\left(\frac{|A \setminus P| \cdot \pi}{2p}\right)\sin\left(\frac{(|A|+|P|-2)\pi}{2p}\right) +$$

$$+ 2\sin\left(\frac{(|P|-|A|)\pi}{2p}\right)\cos\left(\frac{(|A|+|P|-2)\pi}{2p}\right).$$

Now we are almost ready to estimate the size $|A \setminus P|$. First of all, since $|P| \geqslant |A|$, the last summand is positive and can be discarded, leaving us with the inequality

$$\mathbb{E}_x(1_P - 1_A)(x) + 4\varepsilon + 2\lambda + \frac{1}{\mathcal{F}(M)} \geqslant \frac{4\sin\left(\frac{|P \setminus A| \cdot \pi}{2p}\right)\sin\left(\frac{|A \setminus P| \cdot \pi}{2p}\right)\sin\left(\frac{(|A|+|P|-2)\pi}{2p}\right)}{p\sin\left(\frac{\pi}{p}\right)}.$$

If only $4\varepsilon + 2\lambda + \frac{1}{\mathcal{F}(M)} \leqslant (1-\eta)\delta\alpha$ and $|P| \leqslant (1 + (1+\eta)\delta)\alpha p$, we have that the left hand side is bounded by $2\delta\alpha$. On the other hand, if we had $|A \setminus P| > C(\delta\alpha)^{1/2}p$, then the same would hold for $|P \setminus A|$. Knowing the behaviour of sine around 0, we would argue that the first two factors in the numerator are at least $C'(\delta\alpha)^{1/2}$ for some other constant C'. But the last factor is bounded away from 0 (as $|A|$ and $|P|$ are bounded away from both 0 and $\frac{p}{2}$) and the denominator is around π (w.l.o.g. > 3), so this would contradict our inequality. In the end we need to have $|A \setminus P| \leqslant C(\alpha\delta)^{1/2}p$.

The Theorem is now proven up to checking that we can choose all the constants to make the calculations work. First of all, we would like to use the Serra-Zemor Theorem for the set C. We had $|C + C| \leqslant ((2 + \delta)\alpha + \frac{2}{t(\mathcal{F}(M))^{1/2}})p$ and $|C| \geqslant (\alpha - \lambda - \varepsilon)p$. To make sure that $|C + C| < (2 + 10^{-4})|C|$ we want to require $\lambda, \varepsilon < 10^{-6}\alpha$, $\delta < 10^{-6}$ and $t(\mathcal{F}(M))^{1/2} > 10^6\alpha^{-1}$ (say). Then, we would like to have $|C + C| - |C| + 1 = |P| \leqslant (1 + (1+\eta)\delta)\alpha p$. This rearranges to

$$\frac{2}{t(\mathcal{F}(M))^{1/2}} + \varepsilon + \lambda < \eta\delta\alpha.$$

For this it would be enough if $\lambda, \varepsilon, \frac{1}{t(\mathcal{F}(M))^{1/2}} \leqslant \frac{\eta \delta \alpha}{4}$. Moreover, we need $t \leqslant (\frac{\lambda}{2M})^M (\frac{\lambda^2}{4} - 2\sqrt{\varepsilon} - \varepsilon)$. This suggests setting $\varepsilon = \frac{\lambda^4}{256}$ and requiring $t \leqslant (\frac{\lambda}{2M})^M \cdot \frac{\lambda^2}{16}$. We have just listed all the requirements and now the strategy is as follows. Set $\lambda = \frac{\eta \delta \alpha}{4}$ (we can freely assume $\eta < 10^{-6}$ to make sure that $\lambda < 10^{-6}\alpha$) and $\varepsilon = \frac{\lambda^4}{256}$. Now the only thing is to make sure that $\frac{4}{\eta \delta \alpha (\mathcal{F}(M))^{1/2}} \leqslant t \leqslant \frac{\lambda^2}{16}(\frac{\lambda}{2M})^M$. This might seem impossible, as the upper bound on M depends of $\mathcal{F}$ and we might not be able to fit into the correct range. The solution to this problem is the following: suppose that the above inequalities hold for some other number t'. Then the entire argument is correct assuming that the initial inequality describing popular doubling of A holds with parameter t' instead of t. A similar argument to [Maz15, Corollary 3.5] shows that this is indeed the case for any $t' \geqslant t$. This suggest the following strategy:

- given α_1, α_2, η, choose $\delta_0 > 0$ so that $(1 + (1 + \eta)\delta_0)\alpha_2 < \frac{1}{2}$ (to make sure all the hyptheses of Serra-Zemor Theorem are satisfied),
- given $\delta \in (0, \delta_0)$, set $\lambda = \frac{\eta \delta \alpha_1}{4}$ and $\varepsilon = \frac{\lambda^4}{256}$,
- define $\mathcal{F}(M) = \frac{2^{12}}{(\eta \delta \alpha_1 \lambda^2)^2}(\frac{2M}{\lambda})^{2M}$ and apply the arithmetic regularity lemma to get an upper bound $M \leqslant M_0$,
- set $t_0 = \frac{\lambda^2}{16}(\frac{\lambda}{2M_0})^{M_0}$.

Then we can find $t' \geqslant t_0$ with the postulated properties. Since $t \leqslant t_0$, we also have $t \leqslant t'$, as required. That ends the proof of Theorem 1.1.

Appendix A. Arithmetic regularity lemma

In the appendix we give a self-contained proof of the arithmetic regularity lemma for U^2 norm. The lemma was proven in full generality (i.e. for U^k norm for arbitrary k) by Green and Tao in [GT10]. There is also an exposition by Eberhard of the U^2 case. Unfortunately both of them have a feature that is a disadvantage for us, namely they deal with functions defined on $\{1, \ldots, N\}$ rather than $\mathbb{Z}/p\mathbb{Z}$. As a result the "structured part" obtained there comes from a Lipschitz function defined on $[0, 1] \times \mathbb{Z}/q\mathbb{Z} \times \mathbb{T}^d$ (in the U^2 case). However, as we work over a cyclic group of prime order, in our setting everything is periodic $\pmod{p}$ and there is no room either for non-periodic behaviour (such as $[0, 1]$), or periodic behaviour modulo other numbers. Therefore we are aiming for a slightly different statement of the regularity lemma, but the methods of proof remain the same.

Let us by an *pre-character* on a group G mean a homomorphism $\varphi : G \to \mathbb{T}$ and by a *character* a homomorphism $\chi : G \to \{z \in \mathbb{C} \ : \ |z| = 1\}$. Of course there is one-to-one correspondence between those, given by the equation $\chi = e^{2\pi i \varphi}$. This terminology is by no means standard and is used only for the purpose of this paper.

Before we start, let us fix some notation. For any set Γ of pre-characters on $\mathbb{Z}/p\mathbb{Z}$ and any positive integer n, we define a partition $\mathcal{B} = \mathcal{B}(\Gamma, n)$ of $\mathbb{Z}/p\mathbb{Z}$ into cells. Intuitively, $\mathcal{B}$ corresponds to the partition of the torus $\mathbb{T}^\Gamma$ into $n^{|\Gamma|}$ cubes of side length $\frac{1}{n}$. More formally, two points $x, y \in \mathbb{Z}/p\mathbb{Z}$ belong to the same cell if $\varphi(x), \varphi(y) \in [\frac{k_\varphi}{n}, \frac{k_\varphi+1}{n}) \subset \mathbb{T}$ for some $k_\varphi \in \mathbb{Z}$ for all $\varphi \in \Gamma$. Note that if $\Gamma \subset \Gamma'$ and $n|n'$, then $\mathcal{B}(\Gamma', n')$ is a refinement of $\mathcal{B}(\Gamma, n)$, i.e. each cell of the former is a union of cells of the latter.

For any function $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ and any partition $\mathcal{B} = \mathcal{B}(\Gamma, n)$ define conditional expectation $\mathbb{E}(f|\mathcal{B})$ in the standard way, i.e. $\mathbb{E}(f|\mathcal{B})(x)$ is the average of f on the cell of $\mathcal{B}$ containing x. In other words, $\mathbb{E}(f|\mathcal{B})$ is just the orthogonal projection of f onto the space of all $\mathcal{B}$-measurable functions (constant on every cell of $\mathcal{B}$). Note that if $\mathcal{B}'$ is a refinement of $\mathcal{B}$ then $\mathbb{E}(\mathbb{E}(f|\mathcal{B}')|\mathcal{B}) = \mathbb{E}(f|\mathcal{B})$ and more generally $\mathbb{E}(f \cdot \mathbb{E}(g|\mathcal{B})|\mathcal{B}') = \mathbb{E}(f|\mathcal{B}')\mathbb{E}(g|\mathcal{B})$.

Lemma A.1. *Let $n > 0$ and let $\varphi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}$ be an pre-character and let $\chi = e^{2\pi i \varphi}$. Let Γ be a set of characters containing φ and let $\mathcal{B} = \mathcal{B}(\Gamma, n)$. Suppose that $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ is a function with $\|f\|_\infty \leqslant 1$. Then*

$$|\langle f - \mathbb{E}(f|\mathcal{B}), \chi\rangle| \leqslant \frac{2\pi}{n}.$$

Proof. The key idea is that χ is almost constant on each cell of $\mathcal{B}$. More precisely, by the properties of orthogonal projections we have

$$\langle f - \mathbb{E}(f|\mathcal{B}), \chi\rangle = \langle f, \chi - \mathbb{E}(\chi|\mathcal{B})\rangle.$$

But since $\varphi \in \Gamma$, the function $\chi - \mathbb{E}(\chi|\mathcal{B})$ is bounded pointwise by $|1 - e^{2\pi i/n}| \leqslant \frac{2\pi}{n}$ and the claim follows. $\qquad\square$

Corollary A.2. Let $\delta > 0$ and let $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ with $\|f\|_\infty \leqslant 1$. Then there exists a set Γ of pre-characters of size $|\Gamma| \leqslant \frac{4}{\delta^2}$ and $n \leqslant \frac{16}{\delta}$ such that for $\mathcal{B} = \mathcal{B}(\Gamma, n)$ and any character χ we have

$$(1) \qquad\qquad\qquad |\langle f - \mathbb{E}(f|\mathcal{B}), \chi\rangle| \leqslant \delta.$$

Proof. Define $n = \lceil \frac{4\pi}{\delta} \rceil$ and build the set Γ iteratively, at the beginning $\Gamma = \emptyset$. At each stage we ask if the inequality (1) is satisfied for every character. If so, we finish our procedure, otherwise we take a character χ for which the inequality fails and add the corresponding pre-character φ to Γ. Let $\mathcal{B} = \mathcal{B}(\Gamma, n)$ and $\mathcal{B}' = \mathcal{B}(\Gamma \cup \{\varphi\}, n)$. By the previous lemma we know that

$$|\langle f - \mathbb{E}(f|\mathcal{B}'), \chi\rangle| \leqslant \frac{2\pi}{n} \leqslant \frac{\delta}{2}.$$

Combining this with the initial assumption on χ and triangle inequality gives

$$|\langle \mathbb{E}(f|\mathcal{B}') - \mathbb{E}(f|\mathcal{B}), \chi\rangle| \geqslant \frac{\delta}{2}.$$

Now we use Cauchy-Schwarz and the fact that $\mathbb{E}(f|\mathcal{B})$ is an orthogonal projection of $\mathbb{E}(f|\mathcal{B}')$ (as $\mathcal{B}'$ is a refinement of $\mathcal{B}$):

$$\|\mathbb{E}(f|\mathcal{B}')\|_2^2 - \|\mathbb{E}(f|\mathcal{B})\|_2^2 = \|\mathbb{E}(f|\mathcal{B}') - \mathbb{E}(f|\mathcal{B})\|_2^2 \geqslant |\langle \mathbb{E}(f|\mathcal{B}') - \mathbb{E}(f|\mathcal{B}), \chi\rangle|^2 \geqslant \frac{\delta^2}{4}.$$

In other words, it means that adding to Γ the pre-character corresponding to χ increases the value of $\|\mathbb{E}(f|\mathcal{B})\|_2^2$ by at least $\frac{\delta^2}{4}$. Since this quantity can only take values between 0 and 1, this procedure must terminate in at most $\frac{4}{\delta^2}$ steps. In the end we get a set Γ of size at most $\frac{4}{\delta^2}$ satisfying the inequality (1) for each character χ.

It remains to check that the bound on n is correct. For that we can freely assume $\delta \leqslant 1$, which implies $\frac{4\pi}{\delta} \geqslant 4\pi > \frac{25}{2}$. However, for any such number x we have the bound $\lceil x \rceil \leqslant \frac{14}{13}x \leqslant \frac{16}{4\pi}x$. $\qquad\square$

The corollary above says that we can get rid of any large Fourier coefficients using only projections of bounded complexity. The heart of the arithmetic regularity lemma is to iterate this argument. Before we do that, le us explain what a growth function is. A *growth function* is simply an increasing function $\mathcal{F} : (0, +\infty) \to (0, +\infty)$, typically describing how large we need one parameter to be in terms of another parameter. In most applications one can think of $\mathcal{F}$ as of an exponential function $x \mapsto C_1 e^{C_2 x}$.

Proposition A.3 (arithmetic regularity lemma: baby version)**.** *Let $\varepsilon > 0$ and let $\mathcal{F}$ be a growth function. Then there exixts a real number $M_0 = M_0(\varepsilon, \mathcal{F}) > 0$ for*

which the following statement is true. Let p be a prime and $f : \mathbb{Z}/p\mathbb{Z} \to [0,1]$ be a function. Then there exists a number $0 < M \leqslant M_0$ and a decomposition

$$f = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

satisfying the following properties:

- $f_{\mathrm{str}} = \mathbb{E}(f|\mathcal{B})$, where $\mathcal{B} = \mathcal{B}(\Gamma, n)$ for some set Γ of pre-characters and some positive integer n with $|\Gamma|, n \leqslant M$ (f_{str} is structured),
- $\|f_{\mathrm{sml}}\|_2 \leqslant \varepsilon$ (f_{sml} is small),
- $|\langle f_{\mathrm{unf}}, \chi \rangle| \leqslant \frac{1}{\mathcal{F}(M)}$ for every character χ (f_{unf} is U^2-uniform).
- f_{str} and $f_{\mathrm{str}} + f_{\mathrm{sml}}$ both take values in $[0,1]$.

Proof. We will again use an iterative procedure. At the beginning, let $\mathcal{B}$ be the trivial partition, corresponding to $\Gamma = \emptyset$ and $n = 1$. At each stage, we set $M = \max(|\Gamma|, n)$ and then apply Corollary A.2 with parameter $\delta = \frac{1}{\mathcal{F}(M)}$ to the function $f - \mathbb{E}(f|\mathcal{B})$. This way we get a set Γ' and an integer n', both bounded in terms of M and $\mathcal{F}$. In fact, we need a slightly modified version of this result; to ensure that $\Gamma \subset \Gamma'$ and $n|n'$, we take at the beginning $n \cdot \lceil \frac{4\pi}{\delta} \rceil$ instead of $\lceil \frac{4\pi}{\delta} \rceil$ and Γ istead of the empty set. This does not affect the boundedness of the final parameters, we still have the bounds of the shape $|\Gamma'|, n \leqslant \mathcal{F}'(M)$ for some growth function $\mathcal{F}'$ depending only on $\mathcal{F}$. After applying this procedure we wish to set $f_{\mathrm{str}} = \mathbb{E}(f|\mathcal{B})$, $f_{\mathrm{sml}} = \mathbb{E}(f|\ mathcalB)' - \mathbb{E}(f|\mathcal{B})$ and $f_{\mathrm{unf}} = f - \mathbb{E}(f|\mathcal{B})'$, where $\mathcal{B}' = \mathcal{B}(\Gamma', n')$. All the required conditions are clearly satisfied except one: it might happen that $\|f_{\mathrm{sml}}\|_2 > \varepsilon$. To take care of it, we use iteration: if this actually happened, set new $\Gamma := \Gamma'$ and $n := n'$. Again, we can argue that since $\mathcal{B}'$ is a refinement of $\mathcal{B}$, we have

$$\|\mathbb{E}(f|\mathcal{B}')\|_2^2 - \|\mathbb{E}(f|\mathcal{B})\|_2^2 = \|\mathbb{E}(f|\mathcal{B}') - \mathbb{E}(f|\mathcal{B})\|_2^2 = \|f_{\mathrm{sml}}\|_2^2 > \varepsilon^2$$

and therefore each iteration increases the value of $\|\mathbb{E}(f|\mathcal{B})\|_2^2$ by at least ε^2, so we cannot have more than $\frac{1}{\varepsilon^2}$ iterations in total. It means that in the end the result is true with $M_0 = \underbrace{\mathcal{F}'(\ldots(\mathcal{F}'(1))\ldots)}_{\lfloor \frac{1}{\varepsilon^2} \rfloor \text{ iterations}}$. $\qquad\square$

The above version of the arithmetic regularity lemma is not quite satisfactory, as we expect to have a slightly different kind of structure for f_{str}. Before we explain, how to fix that, let us exploit some properties of Fejer kernel.

Lemma A.4. *Let d and K be positive integers, let $\chi_j : \mathbb{T}^d \to \mathbb{C}$ $(j = 1, \ldots, d)$ be the basic characters defined as $\chi_j(t) = e^{2\pi i t_j}$ and let $\Phi_K : \mathbb{T} \to [0, +\infty)$ be the Fejer kernel of order K:*

$$\Phi_K(t) = \frac{1}{K^d} \prod_{j=1}^{d} \left| \sum_{k=0}^{K-1} \chi_j^k(t) \right|^2.$$

Then $\int_{\mathbb{T}^d} \Phi_K(t) dt = 1$ and moreover

$$\int_{[-\lambda, \lambda]^d} \Phi_K(t) dt \geqslant 1 - \frac{d}{4K\lambda^2}.$$

Proof. The first assertion is standard; to prove it one only needs to expand Φ_K as the linear combination of characters and observe that the trivial character comes with coefficient 1. To prove the inequality, let us first note that

$$\left| \sum_{k=0}^{K-1} \chi_j^k(t) \right| = \left| \frac{1 - \chi_j^K(t)}{1 - \chi_j(t)} \right| \leqslant \frac{2}{|1 - e^{2\pi i t_j}|} \leqslant \frac{1}{2\|t_j\|}.$$

Therefore if $d = 1$, we have

$$\int_{[-\lambda,\lambda]} \Phi_K(t)dt = 1 - \int_{\|t\|\geqslant\lambda} \Phi_K(t)dt \geqslant 1 - \sup_{\|t\|\geqslant\lambda} \Phi_K(t) = 1 - \frac{1}{4K\lambda^2}$$

. Now for $d > 1$, the d-dimensional Fejer kernel is just a product of d copies of a 1-dimensional one, which gives the bound

$$\int_{[-\lambda,\lambda]^d} \Phi_K(t)dt \geqslant \left(1 - \frac{1}{4K\lambda^2}\right)^d \geqslant 1 - \frac{d}{4K\lambda^2}.$$

$\square$

Before stating the next result let us set a default norm on $\mathbb{T}^d$ to be the maximum norm, i.e. $\|t\| = \max_{1\leqslant j\leqslant d} \|t_j\|$. Consequently, we call a function $F : \mathbb{T}^d \to \mathbb{C}$ M-Lipschitz, if $|F(t_1) - F(t_2)| \leqslant M\|t_1 - t_2\|$ holds for all $t_1, t_2 \in \mathbb{T}$.

Proposition A.5 (arithmetic regularity lemma: intermediate version). *Let $\varepsilon > 0$ and let $\mathcal{F}$ be a growth function. Then there exixts a real number $M_0 = M_0(\varepsilon, \mathcal{F}) > 0$ for which the following statement is true. Let $p > p_0(\varepsilon, \mathcal{F})$ be a prime and $f : \mathbb{Z}/p\mathbb{Z} \to [0, 1]$ be a function. Then there exists a number $0 < M \leqslant M_0$ and a decomposition*

$$f = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

satisfying the following properties:

- $f_{\mathrm{str}} = F \circ \varphi$, *where* $\varphi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^d$ *is a homomorphism with* $d \leqslant M$ *and* $F : \mathbb{T}^d \to [0, 1]$ *is an M-Lipschitz function (f_{str} is structured),*
- $\|f_{\mathrm{sml}}\|_2 \leqslant \varepsilon$ *(f_{sml} is small),*
- $|\langle f_{\mathrm{unf}}, \chi\rangle| \leqslant \frac{1}{\mathcal{F}(M)}$ *for every character χ (f_{unf} is U^2-uniform).*
- f_{str} *and* $f_{\mathrm{str}} + f_{\mathrm{sml}}$ *both take values in $[0, 1]$.*

Proof. First we apply to f the baby version with some different parameters ε' and $\mathcal{F}'$ to be specified later. We get a decomposition $f = f'_{\mathrm{str}} + f'_{\mathrm{sml}} + f'_{\mathrm{unf}}$; now we set $f_{\mathrm{unf}} = f'_{\mathrm{unf}}$ and try to find f_{str} of the new type so that $\|f_{\mathrm{str}} - f'_{\mathrm{str}}\|_2$ is small and in the end set $f_{\mathrm{sml}} = f'_{\mathrm{sml}} + f'_{\mathrm{str}} - f_{\mathrm{str}}$. We know the structure of f'_{str}; it can be alternatively said that $f'_{\mathrm{str}} = F' \circ \varphi$, where $\varphi = (\varphi_1, \ldots, \varphi_d) : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^d$ is just a product of all pre-characters forming Γ and $F' : \mathbb{T}^d \to [0, 1]$ is a function that is constant on the cubes of the form $[\frac{k_1}{n}, \frac{k_1+1}{n}] \times \ldots \times [\frac{k_d}{n}, \frac{k_d+1}{n}]$ for $k_1, \ldots, k_d \in \mathbb{Z}$. The function F' does not need to be unique, we can pick any that fits into the formula. Also, it does not need to be Lipschitz and we have to fix that. To do this, put $F = F' * \Phi_K$ for some K to be specified. To see that F is Lipschitz, let us calculate

$$|F(t_1) - F(t_2)| = \left|\int_{\mathbb{T}^d} F'(s)(\Phi_K(t_1 - s) - \Phi_K(t_2 - s))\right| \leqslant$$
$$\leqslant \sup_s |\Phi_K(t_1 - s) - \Phi_K(t_2 - s)|.$$

The above calculation shows that the Lipschitz constant of F is bounded by that of Φ_K. To estimate it, let us note that Φ_K is a linear combination of characters of the form $\prod_{j=1}^d \chi_j^{k_j}$ with $|k_j| < K$, and a character of this particular form comes with coefficient $\prod_{j=1}^d (1 - \frac{|k_j|}{K})$ and is itself a $(2\pi \sum_{j=1}^d |k_j|)$-Lipschitz function. Therefore the Lipschitz constant of Φ_K is at most

$$L_d = 2\pi \sum_{k_1,\ldots,k_d} \left(\left(\sum_{j=1}^d |k_j|\right) \cdot \prod_{j=1}^d \left(1 - \frac{|k_j|}{K}\right)\right).$$

To calculate this, let us set $L'_d = \sum_{k_1,\ldots,k_d} \prod_{j=1}^{d} \left(1 - \frac{|k_j|}{K}\right)$. Then one can check that those sequenced satisfy the recurrence $L_{d_1+d_2} = L_{d_1}L'_{d_2} + L_{d_2}L'_{d_1}$ and $L'_{d_1+d_2} = L'_{d_1} + L'_{d_2}$, which together with the boundary conditions $L_1 = \frac{2\pi}{3}(K^2 - 1)$, $L'_1 = K$ gives $L_d = \frac{2\pi}{3}dK^{d-1}(K^2 - 1) \leqslant 4dK^{d+1}$.

Set $f_{\text{str}} = F \circ \varphi$. We would like to bound the expression

$$\|f'_{\text{str}} - f_{\text{str}}\|_2^2 = \mathbb{E}_x |F'(\varphi(x)) - F' * \Phi_K(\varphi(x))|^2.$$

Inside the expectation, some of the elements s will lie near the edges of the cubes and for them it would be hard to estimate the value $|F'(\varphi(x)) - F' * \Phi_K(\varphi(x))|$ other than trivially by 1. Let us estimate the number of such "bad" elements: the set of all $t_j \in \mathbb{T}$ with $\|t_j - \frac{k_j}{n}\| \leqslant \lambda$ has measure 2λ; since p is sufficiently large we can assume that the set of all x with $\|\varphi_j(x) - \frac{k_j}{n}\| \leqslant \lambda$ has size at most $4\lambda p$. Taking into account all possible values of j and k_j we see that all but at most $4\lambda dnp$ elements are separated from the boundary of their cubes by at least λ. For those elements x let us estimate

$$|F'(\varphi(x)) - F' * \Phi_K(\varphi(x))| \leqslant \int_{\mathbb{T}} \Phi_K(t)|F'(\varphi(x)) - F'(\varphi(x) - t)|dt.$$

By the description of x the latter factor is zero on the cube $[-\lambda, \lambda]^d$; on the remaining set it is bounded by 1, so by the previous lemma the value of the integral is bounded by $\frac{d}{4K\lambda^2}$. In the end, taking into account all values of x, we have an estimate

$$\|f'_{\text{str}} - f_{\text{str}}\|_2^2 = \mathbb{E}_x |F'(\varphi(x)) - F' * \Phi_K(\varphi(x))|^2 \leqslant 4\lambda dn + \left(\frac{d}{4K\lambda^2}\right)^2.$$

We wish the last quantity to be at most $\frac{\varepsilon}{2}$; to achieve this set $\lambda = \frac{\varepsilon}{16dn}$ and $K = \lceil \frac{d}{2\lambda^2\sqrt{\varepsilon}} \rceil$.

Now we return to the beginning, where we had to specify ε' and $\mathcal{F}'$. We can take $\varepsilon' = \frac{\varepsilon}{2}$, then $\|f_{\text{sml}}\|_2 \leqslant \|f'_{\text{sml}}\|_2 + \|f'_{\text{str}} - f_{\text{str}}\|_2 \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. To choose $\mathcal{F}'$, let us first note that

$$K \leqslant \frac{d}{\lambda^2\sqrt{\varepsilon}} = \frac{2^8 d^3 n^2}{\varepsilon^{5/2}} \leqslant \frac{2^8 M^5}{\varepsilon^{5/2}}.$$

The Lipschitz constant of F is then bounded by

$$4dK^{d+1} \leqslant 4M \left(\frac{2^8 M^5}{\varepsilon^{5/2}}\right)^{M+1} =: a(M, \varepsilon).$$

It is now enough to take $\mathcal{F}'(M) = \mathcal{F}(a(M, \varepsilon))$ and $M_0 = M_0(\varepsilon', \mathcal{F}')$ given by the previous version of the lemma. $\qquad \square$

Now the structure of f_{str} appears to be more natural, although we are still missing some information. We would like to know that the image of the homomorphism φ is well equidisributed in $\mathbb{T}^d$ so that we could expect that f_{str} has roughly the same global structure as F. To achieve this, let us set a notion of K-independence. The homomorphism $\varphi = (\varphi_1, \ldots, \varphi_d)$ will be called K-independent if the only solution to the equation $k_1\varphi_1 + \ldots + k_d\varphi_d = 0$ with $|k_j| < K$ is $k_1 = \ldots = k_d = 0$. We will show how we can require independence, in particular what to do if φ turns out not to be independent.

Lemma A.6. *Let $(a_1, \ldots a_d)$ be a vector with integer coordinates. There exists a matrix $A = [a_{ij}] \in \mathcal{M}_d(\mathbb{Z})$ with $a_{1j} = a_j$ (for $j = 1, \ldots, d$) and satisfying the*

following properties:

$$\sum_{j=1}^{d} a_j a_{ij} = 0 \qquad \text{for } i = 2, \ldots, d,$$

$$\det A = \frac{\sum_{j=1}^{d} a_j^2}{\gcd(a_1, \ldots, a_d)}.$$

Moreover, if $\max_{1 \leqslant j \leqslant d} |a_j| \leqslant K$, *then we can choose the entries of the matrix A to be bounded by K.*

Proof. Let us prove the claim by induction on d. For $d = 1$ the statement is trivial. Suppose $d > 1$ and we have already proved it for $d-1$. We would like to extend the matrix found for the vector $(a_1, \ldots, a_{d-1})$ to make it work for $(a_1, \ldots, a_d)$. Setting $a_{id} = 0$ for $i = 2, \ldots, d-1$ makes the first property satisfied for those values of i. Also, it makes the determinant quite easy to calculate by expanding it with respect to the last column. Since

$$\frac{\sum_{j=1}^{d} a_j^2}{\gcd(a_1, \ldots, a_d)} = \frac{\sum_{j=1}^{d-1} a_j^2}{\gcd(a_1, \ldots, a_{d-1})} \cdot \frac{\gcd(a_1, \ldots, a_{d-1})}{\gcd(a_1, \ldots, a_d)} + \frac{a_d^2}{\gcd(a_1, \ldots, a_d)},$$

it looks reasonable to set $a_{dd} = \frac{\gcd(a_1, \ldots, a_{d-1})}{\gcd(a_1, \ldots, a_d)} \in \mathbb{Z}$ and try to complete the last row so that the submatrix B obtained by deleting the first row and the last column has determinant $\frac{(-1)^{d-1} a_d}{\gcd(a_1, \ldots, a_d)} \in \mathbb{Z}$. Note that the formula

$$\det B = \frac{(-1)^d \sum_{j=1}^{d-1} a_j a_{dj}}{\gcd(a_1, \ldots, a_{d-1})}$$

is true if we set $a_{dj} = a_{ij}$ for some $1 \leqslant i \leqslant d-1$ and all $j = 1, \ldots d-1$. Since the vectors $(a_{ij})_{j=1}^{d-1}$ span all of $\mathbb{R}^{d-1}$ (as the determinant of the matrix they form is non-zero by the inductive hypothesis), the formula above is in fact true for any choice of $a_{d,1}, \ldots, a_{d,d-1}$. This is good for us — if we insist that

$$0 = \sum_{j=1}^{d} a_j a_{dj} = \sum_{j=1}^{d-1} a_j a_{dj} + \frac{a_d \gcd(a_1, \ldots, a_{d-1})}{\gcd(a_1, \ldots, a_d)},$$

then automatically we have

$$\det A = a_{dd} \cdot \frac{\sum_{j=1}^{d-1} a_j^2}{\gcd(a_1, \ldots, a_{d-1})} + (-1)^{d-1} a_d \cdot \det B =$$

$$\frac{\sum_{j=1}^{d-1} a_j^2}{\gcd(a_1, \ldots, a_d)} - \frac{a_d \sum_{j=1}^{d-1} a_j a_{dj}}{\gcd(a_1, \ldots, a_{d-1})} = \frac{\sum_{j=1}^{d} a_j^2}{\gcd(a_1, \ldots, a_d)}.$$

So the only condition remaining is $\sum_{j=1}^{d-1} a_j a_{dj} = \frac{-a_d}{\gcd(a_1, \ldots, a_d)} \cdot \gcd(a_1, \ldots, a_{d-1})$. This can be satisifed by the Euclidean algorithm since $\frac{-a_d}{\gcd(a_1, \ldots, a_d)} \in \mathbb{Z}$, and thus we have proved the existence of the matrix A.

Now let us consider the bounds for the entries. Obviously if $|a_j| \leqslant K$, then $|a_{dd}| = \left| \frac{\gcd(a_1, \ldots, a_{d-1})}{\gcd(a_1, \ldots, a_d)} \right| \leqslant K$. Choosing the vector $(a_{d,1}, \ldots, a_{d,d-1})$ carefully might be a little bit more complicated. But if we write the equation in the form

$$\sum_{j=1}^{d-1} a_{dj} \cdot \frac{a_j}{\gcd(a_1, \ldots, a_{d-1})} = -\frac{a_d}{\gcd(a_1, \ldots, a_d)},$$

we can see that the claim boils down to the lemma below.   □

Lemma A.7. *Let $m, K > 0$ be integers and let $b_1, \ldots, b_m$ be coprime integers not exceeding K in absolute value. Let b be an integer with $|b| \leqslant K$. Then there exist integers $c_1, \ldots, c_m$ not exceeding K in absolutee value and satisfying*

$$b_1 c_1 + \ldots + b_m c_m = b.$$

Proof. If $m = 1$, then $b_1 = \pm 1$ and thee statement is trivial. For $m = 2$, if either of b_1, b_2 is equal to ± 1, the statement is trivial as well. If it is not the case, then without loss of generality we can assume $b_1 > b_2 > 0$. But then the numbers $b - k b_2$ for $|k| \leqslant K$ are at most $K b_1$ in absolute value and at least one of them is a multiple of b_1, so the claim follows. Suppose now that $m > 2$ and we have already proven the claim for all smaller values of m. Let $g = \gcd(b_1, \ldots, b_{m-1})$. Then b_m and g are coprime and have absolute value at most K, so we can use the claim for $m = 2$ to get $g b' + b_m c_m = b$ with $|b'|, |c_m| \leqslant K$. Also the numbers $\frac{b_1}{g}, \ldots, \frac{b_{m-1}}{g}$ are coprime integers bounded by K in absolute value, which gives us $\frac{b_1}{g} c_1 + \ldots + \frac{b_{m-1}}{g} c_{m-1} = b'$ with $|c_1|, \ldots, |c_{m-1}| \leqslant K$. Combining those two identities we get the claim. $\square$

Lemma A.6 gives us the following corollary.

Corollary A.8. *Let $d > 1$ be an integer, let $\varphi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^d$ be a homomorphism and let $F : \mathbb{T}^d \to \mathbb{C}$ be an M-Lipschitz function. Then at least one of the following holds:*

- *φ is K-independent,*
- *there exits a homomorphism $\varphi' : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^{d-1}$ and a dKM-Lipschitz function $F' : \mathbb{T}^{d-1} \to \mathbb{C}$ with $F' \circ \varphi' = F \circ \varphi$.*

Proof. Suppose φ is not K-independent, i.e. there exist intgers $a_1, \ldots, a_d$ with $\sum_{j=1}^{d} a_j \varphi_j = 0$ and $|a_j| < K$. It is not hard to see that the second part is true for $p \leqslant K$ as long as $d - 1 \geqslant 1$; suppose then $p > K$. In that case we are allowed to divide all of a_j by their greatest common divisor and without loss of generality assume $\gcd(a_1, \ldots, a_d) = 1$. By Lemma A.6 we can find $d - 1$ integer vectors orthogonal to $a = (a_1, \ldots, a_d)$ suth that the matrix A consisting of all of them has determinant $\sum_{j=1}^{d} a_j^2$. We claim that the $\mathbb{Z}$-span of these $d - 1$ vectors coincides with the intersection of their $\mathbb{R}$-span and $\mathbb{Z}^d$. Indeed, we know that the $\mathbb{Z}$-span of all d vectors is a subgroup of $\mathbb{Z}^d$ of index $\det A = \sum_{j=1}^{d} a_j^2$. On the other hand, the map $x \mapsto \langle a, x \rangle \pmod{\det A}$ gives rise to a surjective homomorphism from the quotient group to a group of size $\det A$. Therefore this homomorphism is in fact an isomorphism and its kernel is precisely the $\mathbb{Z}$-span of all d vectors. Intersecting it with the $\mathbb{R}$-span of the last $d - 1$ vecrors obviously gives us their $\mathbb{Z}$-span. But this intersection can be easily seen as $\{x \in \mathbb{Z}^d : \langle a, x \rangle = 0\}$ or in other words the intersection of the $\mathbb{R}$-span and all of $\mathbb{Z}^d$.

Since $\gcd(a_1, \ldots, a_d) = 1$, it follows that the set $\{t \in \mathbb{T}^d : \langle a, t \rangle = 0 \in \mathbb{T}\}$ is in fact the image of the subspace $\{x \in \mathbb{R}^d : \langle a, x \rangle = 0 \in \mathbb{R}\}$ under the projection $\pmod{\mathbb{Z}^d}$. Therefore it can be parametrized as $A'(\mathbb{T}^{d-1})$, where A' is $(d-1) \times d$ matrix obtained from A by deleting its first row $(a_1, \ldots, a_n)$. Let $F' : \mathbb{T}^{d-1} \to \mathbb{C}$ and $\varphi' : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^{d-1}$ be functions satisfying $F' = F \circ A'$ and $\varphi = A' \circ \varphi'$. Note that φ' is well defined and is a homomorphism. Then

$$F' \circ \varphi' = F \circ A' \circ \varphi' = F \circ \varphi.$$

The only thing left is to estimate the Lipschitz constant of F'. Since A' has entries bounded by K, it can be viewed as a dK-Lipshitz function. Composing it with an M-Lipschitz function F gives us a function of Lipschitz constant at most dKM. $\square$

Now we are ready to give a proof of the full version of the regularity lemma (in the U^2 case).

Theorem A.9 (arithmetic regularity lemma: final version)**.** *Let $\varepsilon > 0$ and let $\mathcal{F}$ be a growth function. Then there exixts a real number $M_0 = M_0(\varepsilon, \mathcal{F}) > 0$ for which the following statement is true. Let $p > p_0(\varepsilon, \mathcal{F})$ be a prime and $f : \mathbb{Z}/p\mathbb{Z} \to [0,1]$ be a function. Then there exists a number $0 < M \leqslant M_0$ and a decomposition*

$$f = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

satisfying the following properties:

- *$f_{\mathrm{str}} = F \circ \varphi$, where $\varphi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^d$ is a $\mathcal{F}(M)$-independent homomorphism with $d \leqslant M$ and $F : \mathbb{T}^d \to [0,1]$ is an M-Lipschitz function (f_{str} is structured),*
- *$\|f_{\mathrm{sml}}\|_2 \leqslant \varepsilon$ (f_{sml} is small),*
- *$|\langle f_{\mathrm{unf}}, \chi \rangle| \leqslant \frac{1}{\mathcal{F}(M)}$ for every character χ (f_{unf} is U^2-uniform).*
- *f_{str} and $f_{\mathrm{str}} + f_{\mathrm{sml}}$ both take values in $[0,1]$.*

Proof. Let us start with applying the previous version of the regularity lemma with the same parameter ε and a different growth function $\mathcal{F}'$ to be specified. We stick to the obtained decomposition $f = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$ but we would like to exploit more properties of F and φ. If φ is $\mathcal{F}$-independent, we are done. Otherwise we use Lemma A.8 to decrease the dimension d by 1 at the cost of potentially increasing the Lipschitz constant up to $dM\mathcal{F}(M) \leqslant M^2\mathcal{F}(M)$. Put $\mathcal{F}_1(M) = M^2\mathcal{F}(M)$; since this procedure can be applied at most $d \leqslant M$ times, so the correct choice of $\mathcal{F}'$ is $\mathcal{F}'(M) = \mathcal{F}(\underbrace{\mathcal{F}_1(\ldots(\mathcal{F}_1(M))\ldots))}_{\lfloor M \rfloor \text{ times}}$. $\square$

To make a proper use of the above result, we often need to relate the behaviours of f_{str} and F. Below we prove a statement of this kind.

Lemma A.10. *Let d be a positive integer, p be a prime, let $\varphi_1, \ldots, \varphi_d : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}$ be pre-characters and let $\varphi = (\varphi_1, \ldots, \varphi_d) : \mathbb{Z}/p\mathbb{Z} \to \mathbb{T}^d$ be their product. Suppose that the set $\{\varphi_j\}_{j=1}^d$ is K-independent. Let $F : \mathbb{T}^d \to \mathbb{C}$ be an M-Lipschitz function. Then*

$$\left| \mathbb{E}_x F(\varphi(x)) - \int_{\mathbb{T}^d} F(t)dt \right| \leqslant \frac{M}{\sqrt{K}}.$$

Proof. Let $\chi_j : \mathbb{T}^d \to \mathbb{C}$ defined via $\chi(t_1 \ldots, t_d) = e^{2\pi i t_j}$ be the basic characters and let $\Phi_K : \mathbb{T} \to \mathbb{C}$ be the Fejer kernel defined via the formula

$$\Phi_K(t) = \frac{1}{K^d} \prod_{j=1}^d \left| \sum_{k=0}^{K-1} \chi_j^k(t) \right|^2.$$

Recall that $\int_{\mathbb{T}^d} \Phi_K(t)dt = 1$. We also have the bound

$$\left| \sum_{k=0}^{K-1} \chi_j^k(t) \right| = \left| \frac{1 - \chi_j^K(t)}{1 - \chi_j(t)} \right| \leqslant \frac{2}{|1 - e^{2\pi i t_j}|} \leqslant \frac{1}{2\|t_j\|}.$$

Therefore $\Phi_K(t) \leqslant \prod_{j=1}^d \frac{1}{4K\|t_j\|^2}$. Having this inequality we would like to show that $F(t)$ and $F * \Phi_K(t)$ are close together for any $t \in \mathbb{T}^d$. Let us estimate their difference:

$$|F(t) - F * \Phi_K(t)| = \left| \int_{\mathbb{T}^d} (F(t) - F(t-s))\Phi_K(s)ds \right| \leqslant M \int_{\mathbb{T}^d} \|s\|\Phi_K(s)ds.$$

This is already independent of t; now we make use of the fact that $\Phi_K(s)$ is large precisely when $\|s\|$ is small; more accurately $\|s\| \geqslant u$ implies $\Phi_K(s) \leqslant (\frac{1}{4Ku^2})^d$ for

any $u \geqslant 0$. Combining this with $\int_{\mathbb{T}^d} \Phi_K(t)dt = 1$ leads to the inequality

$$\int_{\mathbb{T}^d} \|s\| \Phi_K(s) ds = \int_{\mathbb{T}^d} \int_0^{1/2} 1_{\{u \leqslant \|s\|\}} \Phi_K(s) du ds =$$

$$= \int_0^{1/2} \int_{\mathbb{T}^d} 1_{\{u \leqslant \|s\|\}} \Phi_K(s) ds du \leqslant \int_0^{1/2} \min\left(1, \frac{1}{4Ku^2}\right)^d du \leqslant \frac{1}{\sqrt{K}},$$

and consequently $|F(t) - F * \Phi_K(t)| \leqslant \frac{M}{\sqrt{K}}$ for any $t \in \mathbb{T}^d$. In particular, the triangle inequality yields $|\mathbb{E}_x F(\varphi(x)) - \mathbb{E}_x F * \Phi_K(\varphi(x))| \leqslant \frac{M}{\sqrt{K}}$. Let us now expand the last expression:

$$\mathbb{E}_x F * \Phi_K(\varphi(x)) = \mathbb{E}_x \left(\int_{\mathbb{T}^d} F(t) \Phi_K(\varphi(x) - t) dt \right) =$$

$$= \int_{\mathbb{T}^d} F(t) \left(\mathbb{E}_x \Phi_K(\varphi(x) - t) \right) . dt$$

Expanding out the formula for Φ_K, we can see that the average $\mathbb{E}_x \Phi_K(\varphi(x) - t)$ is a sum of the averages of the form $c \cdot \mathbb{E}_x \chi(\varphi(x) - t)$, where $\chi(t) = \prod_{j=1}^d \chi_j^{\alpha_j}$ for some $\alpha_1, \ldots \alpha_j \in \{1 - K, \ldots, -1, 0, 1, \ldots, K - 1\}$. But from the K-independence of φ we can see that $\chi(\varphi(x) - t)$ is never a constant function in x and therefore has average 0, unless χ is a trivial chatracter with $\alpha_1 = \ldots = \alpha_j = 0$, in which case $c = 1$. In the end, $\mathbb{E}_x \Phi_K(\varphi(x) - t) = 1$ for all $t \in \mathbb{T}^d$, which leads to $\mathbb{E}_x F * \Phi_K(\varphi(x)) = \int_{\mathbb{T}^d} F(t) dt$. Plugging this formula into the previous inequality, we get the desired result. $\square$

One may wonder how we managed to prove the above lemma without assuming that p is large. In fact, the K-independence of a set of d homomorphisms carries a hidden assumption $p \geqslant K^d$.

In this paper we did not need the above result, or even the full version of the regularity lemma (the intermediate version would be enough), but in general it might be useful to have them around.

REFERENCES

[GT10] Ben Green and Terence Tao. An arithmetic regularity lemma, an associated counting lemma, and applications. In *An irregular mind*, volume 21 of *Bolyai Soc. Math. Stud.*, pages 261–334. János Bolyai Math. Soc., Budapest, 2010

[Maz15] Przemysław Mazur, A structure theorem for sets of small popular doubling, submitted to *Acta Arithmetica*

[SZ09] [27] O. Serra and G. Z'emor, Large sets with smaU doubhng modulo p are well covered by arithmetic progressions, *Ann. L'Institut Fomier*, 59 (5) (2009),

[TV06] T. Tao, V. Vu, Additive combinatorics, *Cambridge University Press* 2006

MATHEMATICAL INSTITUTE, RADCLIFFE OBSERVATORY QUARTER, WOODSTOCK ROAD, OXFORD OX2 6GG, UNITED KINGDOM

E-mail address: `przemyslaw.mazur@maths.ox.ac.uk`